Mon. Not. R. Astron. Soc. **000**, 1–6 (2006) Printed 17 November 2006 (MN LATEX style file v2.2)

# The Axis of Evil revisited

Kate Land[1] and João Magueijo[2,3,4]
[1] *Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*
[2] *Perimeter Institute for Theoretical Physics, 31 Caroline St N, Waterloo N2L 2Y5, Canada*
[3] *Canadian Institute for Theoretical Astrophysics, 60 St George St, Toronto M5S 3H8, Canada*
[4] *Theoretical Physics Group, Imperial College, Prince Consort Road, London SW7 2BZ, UK*
*Email contact: krl@astro.ox.ac.uk, jmagueijo@perimeterinstitute.ca*



Accepted xxx. Received xxx; in original form xxx

**ABSTRACT**
In light of the three-year data release from WMAP we re-examine the evidence for the "Axis of Evil". We discover that previous frequentist methods are not robust with respect to the data-sets available and different treatments of the galactic plane. We identify the cause of the instability and show that this result is not a weakness of the data. This is further confirmed by exhibiting an alternative approach, Bayesian in flavour, and based on a likelihood method and the information criteria. We find strong (and sometimes decisive) evidence for the "Axis of Evil" in almost all renditions of the WMAP data. However some significant differences between data-sets remain, and the quantitative aspects of the result depend on the particular information criteria used.

**Key words:** cosmic microwave background

## 1 INTRODUCTION

The Wilkinson Microwave Anisotropy Probe (WMAP) has produced spectacular high resolution all-sky observations of the Cosmic Microwave Background (CMB), which have bolstered the case for the ΛCDM concordance cosmological model (Spergel et al. 2003, 2006). After the release of the first-year results (Bennett et al. 2003) there was a flurry of studies into the Gaussianity and statistical isotropy of the data, as these are fundamental predictions of inflation theories. Reports of something awry have been obtained using a variety of techniques (*e.g.*, Park (2004); Eriksen et al. (2004a); Hansen et al. (2004); Donoghue & Donoghue (2005); Land & Magueijo (2005a); Hansen et al. (2004); Eriksen et al. (2004b); Vielva et al. (2004)). In this paper we focus on anomalies in the largest scale modes, after it was first noted that the quadrupole ($\ell = 2$) and octopole ($\ell = 3$) appeared to be correlated (de Oliveira-Costa et al. 2004), and their power is suspiciously low. Much work has focussed on the alignment and "planarity" of these two multipoles (Copi et al. 2006; Schwarz et al. 2004; Ralston & Jain 2004); but in Land & Magueijo (2005b) it was seen that the alignment actually extends to the *four* multipoles $\ell = 2-5$, along the axis $(b, l) \approx (60, -100)$. This feature has been dubbed the "axis of evil" (AOE).

To be more precise the AOE expression has come to signify various different things. Generally it is intended to denote any form of statistical anisotropy, *i.e.*, a feature in the CMB fluctuations which picks a preferred direction. This can be realized in many ways *e.g.*, multipole planarity (the dominance of $m = \pm\ell$ modes along the preferred axis), or a more general form of $m$-preference. In this respect it must be said that while everyone agrees on the presence of the "axis of evil" *in the data*, its extent is still debated. The expression is also sometimes associated with the low power in the low $\ell$s. This is quite inappropriate: while low power may be related to the AOE (see Land & Magueijo (2006)) there is nothing "axial" or anisotropic in a power spectrum anomaly per se.

There are two possible fault lines in the analysis leading to the "axis of evil" effect. The first concerns the integrity of the data itself, *i.e.*, contamination from noise, systematics and foregrounds. Comparison between the first-year (WMAP1) and third-year (WMAP3) data releases shows that the raw data has hardly changed on large scales. However there are several "all-sky" renditions of the data and these do lead to significant disparities: in this paper we show that this is true regarding the intensity of the AOE, so that discussions should emphasise not so much 1st V's 3rd year data, but the various treatments of the galactic plane region.

The second fault line concerns the "meaning" of the detection, and by this we mean the robustness of the statistics used, and whether there is support for planarity or more general $m$-preference. The frequentist formalism provides no clean way to penalise for extra parameters or to weigh-up the detections against each other, or the null hypothesis. Instead simulations are used to assess how likely it is to get such a feature in a Gaussian SI CMB sky, but selection effects are

hard to account for. Here the confrontation of Bayesianism and frequentism becomes a very practical matter.

We carry out this project as follows. In Section 2 we examine the frequentist AOE results for various renditions of the WMAP1 and WMAP3 data, and we discuss further the limitations of the original frequentist method, such as its lack of robustness, at least with regards to $m$-preference AOE (as opposed to "planarity"). In Section 3 we follow a Bayesian treatment, and find that this is much more robust when confronted with the different data-sets. Further we can compare the evidence for planarity and for more complex $m$-preference. In Section 4 we summarise and discuss the results.

## 2 INSTABILITIES OF THE FREQUENTIST STATISTICS

To assign an axis to each multipole, de Oliveira-Costa et al. (2004) proposed the following statistic:

$$q_\ell = \max_{\mathbf{n}} \left[ \sum_m m^2 |a_{\ell m}(\mathbf{n})|^2 \right], \qquad (1)$$

where the $a_{\ell m}$s are computed in the frame with $\mathbf{z}$-axis in direction $\mathbf{n}$. This selects the frame dominated by the planar $m = \pm\ell$ modes.

In Land & Magueijo (2005b) we generalized this statistic to allow for any $m$ domination, *i.e.*, not restricting ourselves to planar configurations, with the statistic:

$$r_\ell = \max_{m\mathbf{n}} \left[ \frac{C_{\ell m}(\mathbf{n})}{\sum |a_{\ell m'}|^2} \right], \qquad (2)$$

where $C_{\ell 0}(\mathbf{n}) = |a_{\ell 0}|^2$, and $C_{\ell m}(\mathbf{n}) = 2|a_{\ell m}|^2$ for $m > 0$ (notice that 2 modes contribute for $m \neq 0$), for the $a_{\ell m}$s computed in the frame with $\mathbf{z} = \mathbf{n}$. This produces three important quantities for each multipole: the direction $\mathbf{n}_\ell$, the "shape" $m_\ell$, and the ratio $r_\ell$ of the multipole's power absorbed by the mode $m_\ell$ in the direction $\mathbf{n}_\ell$.

We extend the work of Land & Magueijo (2005b) by applying this statistic to the following data-sets:

• The WMAP mission (Bennett et al. 2003) produced full sky CMB maps from ten differencing assemblies (DAs). They also produced an "internal linear combination" (ILC) map. This assumes no external information about the foregrounds and combines smoothed frequency maps with weights chosen to minimize the rms fluctuations, using separate sets of weights for 12 disjoint sky regions. In the first-year data release the WMAP collaboration advised that the ILC map be used only as a visual tool. However, for the third-year release a thorough error analysis of the ILC map was performed, and a bias correction implemented (Hinshaw et al. 2006). The resulting third-year ILC map (herein WMAP3) is expected to be clean enough on scales $\ell \lesssim 10$ to be used without a mask. WMAP data is available from http://lambda.gsfc.nasa.gov.

• Third-party maps include those of Tegmark et al. (2003), who produced their own ILC map. Like above, an "internal" method is employed assuming only a black-body spectrum for the CMB, but now the weights depend on scale (in harmonic space) as well as galactic latitude. This is advantageous because different sources of contamination dominate at different scales - foregrounds at large scales, and noise at smaller scales. As well as the cleaned map, a Wiener filtered map is produced that, through a comparison with the WMAP best estimates of theoretical $C_\ell$, adjusts the power of the map so to suppress noisy fluctuations. We use their first (TOHc1, TOHw1) and third-year (TOHc3, TOHw3) maps, all available from www.hep.upenn.edu/ max/wmap.html.

• In an anlaysis of the ILC map-making method, Eriksen et al. (2004) proposed a faster algorithm for the computation of the weights, that employs Lagrangian multipliers to linearize the problem. Although this produces identical results to that of the WMAP team, and is indeed the method employed by the WMAP collaboration for their third-year map, the authors applied it to the first-year data using slightly different regions, thus producing a slightly different ILC map (herein LILC1), available at www.astro.uio.no/ hke/cmbdata/WMAP_ILC_lagrange.fits.

There are of course the original frequency maps, which require a mask. However, for the task of assessing statistical isotropy we require full sky information, and thus we only employ these ILC maps. Although this process can never be perfect, in the future the community will require that these maps be clean enough for non-Gaussianity tests, as well as the extraction of the power spectrum.

In Table 1 we list the results obtained with frequentist AOE statistic (2) for the various data-sets. It is clear that this statistic is not robust - very similar maps can find very different results as indicated by the final column. The expected inter-angle for isotropic axes is 1 radian ($\sim 57^\circ$), thus a mean of $\sim 22^\circ$ is remarkably low and a comparison to simulations puts this at the 99.9% confidence level (Land & Magueijo 2005b). However, this result only holds for two of the maps, and a small fluctuation in just one multipole makes the $\mathbf{n}_\ell$ jump elsewhere. This highlights one weakness of this statistic - its discontinuous nature.

In Fig. 1 we visualize how "close calls" may arise, explaining the discontinuities of the results in Table 1. For the quadrupole and octopole of the TOHcl1 map, we plot the power ratio at the position $\mathbf{n}$

$$R_\ell(\mathbf{n}) = \max_m \frac{C_{\ell m}(\mathbf{n})}{\sum |a_{\ell m'}|^2}. \qquad (3)$$

Thus the "axis of evil" statistic (2) picks out $\mathbf{n}_\ell$ as the position of the hottest spot from these maps (note the degeneracy between $m = 1, 2$ for $\ell = 2$ - we avoid this in practice by taking just the $m = 2$ solution). Below the $R_\ell$ maps we plot the associated $m$ picked by $R_\ell$ for a given $\mathbf{n}$. We can now diagnose the instabilities in Table 1, by identifying close calls in the competition for the hottest spot. For the quadrupole the $m = 0$ and $m = 2$ modes, and for the octopole the $m = 1$ and $m = 3$ modes are fighting a close battle. The overall mean inter-angle (which measures the strength of the AOE) depends closely on this battle, and thus the instability of this statistic. We should stress that the instabilities identified here do not seem to plague statistics for planarity (Magueijo & Sorkin 2006).

## 3 THE BAYESIAN ASSESSMENT

The instabilities discovered appear to be cured by a Bayesian treatment, which also allows for an evaluation of the evidence for $m$ preference in $\ell = 2-5$, over simple planarity for

| | $\ell=2$ | | $\ell=3$ | | $\ell=4$ | | $\ell=5$ | | Mean |
|---|---|---|---|---|---|---|---|---|---|
| Map | $(b,l)$ | m | $(b,l)$ | m | $(b,l)$ | m | $(b,l)$ | m | inter-$\theta$ |
| LILC1 | (0.9, 156.7) | 0 | (63.0, -126.9) | 3 | (56.7, -163.7) | 2 | (48.6, -94.7) | 3 | 51.4 |
| TOHc1 | (58.5, -102.9) | 2 | (62.1, -120.6) | 3 | (57.6, -163.3) | 2 | (48.6, -93.4) | 3 | 22.4 |
| TOHw1 | (56.7, -103.4) | 2 | (61.2, -120.6) | 3 | (57.6, -163.3) | 2 | (48.6, -93.4) | 3 | 22.3 |
| TOHc3 | (76.5, -134.0) | 2 | (27.0, 51.9) | 1 | (35.1, -130.6) | 1 | (47.7, -94.7) | 3 | 53.8 |
| TOHw3 | (75.6, -133.2) | 2 | (27.0, 51.9) | 1 | (35.1, -130.6) | 1 | (47.7, -93.3) | 3 | 53.9 |
| WMAP3 | (2.7, -26.5) | 0 | (62.1, -122.6) | 3 | (34.2, -131.2) | 1 | (47.7, -96.0) | 3 | 53.7 |

**Table 1.** The $\mathbf{n}_\ell$ axes, in galactic coordinates $(b,l)$, and $m$ that maximise (2) for the multipoles $\ell=2-5$, for various all-sky renditions of the first and third-year WMAP data. Note the low mean inter-angle values for the TOHcl1 and TOHw1 maps, which indicate a strong correlation between the multipoles (it i.e., AOE). The dicontinuous nature of the statistic causes the results to vary widely.

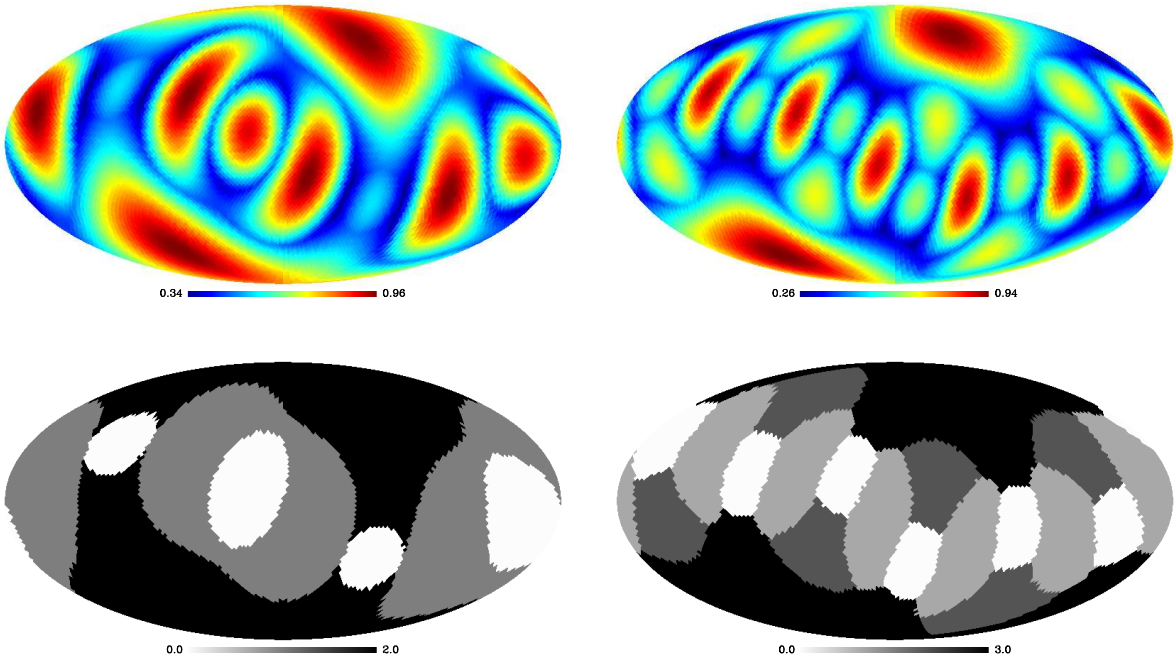


**Figure 1.** The power ratio $R_\ell(\mathbf{n})$ in the dominating $m$ mode (above), and the $m$ value (below) for the quadrupole (left) and octopole (right). The "axis of evil" statistic in (2) searches for the hottest spot in these maps. We can see the close calls that cause the results to vary widely in Table 1. Plotted in galatic coordinates and Mollweide projection.

$\ell=2,3$. Rather than computing a statistic from the maps (*e.g.*, the mean inter-angle between the $\mathbf{n}$ for the various $\ell$), the idea is to assess the "evidence" for a model encoding $m$-preference or planarity. We first outline the general formalism.

Let $\mathcal{L}$ be the likelihood of the data given a model, and $k$ the number of parameters of this model. The parameters should be tuned so to maximize the likelihood, or equivalently, to minimize the information in the data given the theory (defined as $I(D|T)=-\ln(\mathcal{L})$). However the real evidence should refer to the information in the data *and* the theory: $I(D\cap T)=I(D|T)+I(T)$, where the information in the theory, $I(T)$, provides a penalization related to the number of parameters. This matter is behind the "Occam's razor" rationale (Magueijo & Sorkin 2006), and the information criteria (Liddle 2004). There is some controversy on the criterium for $I(T)$, the Aikaike information criterium (AIC) and the Bayesian information criterium (BIC) providing two formalisms that do not always agree. According to the AIC, the information in a theory is simply the number of parameters, $I^{AIC}=k$. The BIC introduces a penalization of $I^{BIC}=\frac{k}{2}\ln N$ where $N$ is the number of data points being fit. We explore both without further comment on their relative merits.

The evidence $H$ for a theory $T_1$ is then defined as the decrease in the information of data and theory when it is compared with a null hypothesis $T_0$:

$$\begin{aligned} H &= I(D\cap T_0)-I(D\cap T_1) \\ &\equiv H_f-H_p, \end{aligned} \tag{4}$$

where $H_f$ measures the improvement in the fit $H_f=\ln(\mathcal{L}_1)-\ln(\mathcal{L}_0)$, and $H_p$ is the extra penalization we have in our new theory. As a rule of thumb $H$ between 3 and 5 sig-

| Data-set | $(b$ | $l)$ | $\epsilon$ | $H_f$ | $H^{AIC}$ | $H^{BIC}$ |
|---|---|---|---|---|---|---|
| LILC1 | 63 | -120 | .042 | 6.51 | 3.51 | 2.78 |
| TOHc1 | 61 | -113 | .032 | 7.48 | 4.48 | 3.75 |
| TOHw1 | 60 | -113 | .032 | 7.32 | 4.32 | 3.59 |
| TOHc3 | 74 | -129 | .018 | 6.97 | 3.97 | 3.24 |
| TOHw3 | 73 | -128 | .018 | 6.94 | 3.94 | 3.21 |
| WMAP3 | 64 | -123 | .043 | 6.49 | 3.49 | 2.76 |

**Table 2.** The raw evidence $H_f$ and the AIC and BIC evidence for the planarity of the $\ell = 2, 3$ modes, using various all-sky renditions of the WMAP data.

| $\ell$ | $m'$ | $(b$ | $l)$ | $\epsilon$ | $H_f$ |
|---|---|---|---|---|---|
| 2 | 0 | 6 | 157 | 0.027 | 3.47 |
| 2 | 2 | 59 | -103 | 0.030 | 3.09 |
| 3 | 3 | 62 | -120 | 0.025 | 5.06 |
| 4 | 2 | 58 | -163 | 0.041 | 5.07 |
| 4 | 0 | 43 | -98 | 0.043 | 4.02 |
| 5 | 3 | 49 | -93 | 0.026 | 7.65 |

**Table 3.** The evidence $H_f$ for a dominating $m$-mode in the TOHcl map, where each multipole can select its own axis, $\epsilon$, and $m'$. Where there is a close call, the runner up $m'$ is also listed.

nals strong evidence, whereas "decisive" evidence requires $H > 5$.

It was shown in Magueijo & Sorkin (2006) that the planarity of the $\ell = 2, 3$ multipoles is supported by a Bayesian analysis. The model used to assess the evidence for planarity is based on the diagonal covariance matrix:

$$\langle |a_{\ell m}|^2 \rangle (\mathbf{n}) = c_\ell \left( \delta_{\ell |m|} + \epsilon (1 - \delta_{\ell |m|}) \right) \qquad (5)$$

where $\mathbf{n}$ and $\epsilon$ are the free parameters of the model (in addition to $c_\ell$ that is common with the isotropic model, but of a different value), with $\epsilon \leqslant 1$. We use the same $\epsilon$ and $\mathbf{n}$ for both multipoles, so that $k = 3$, $N = 12$, $H_p^{AIC} = 3$, and $H_p^{BIC} = 3.73$. In Table 2 we list the parameter values that maximize the likelihood, together with $H_f$ and $H$ following the AIC and BIC methods.

As before (Magueijo & Sorkin 2006) we find that variations between different galactic plane treatments lead to only small variation in $H_f$. They all support strong evidence for planarity using the AIC, but using the stricter BIC our more exhaustive analysis reveals two data-sets where $H$ drops below 3: the LILC1 and the WMAP3.

Using the same formalism we now revisit the debate on the extent of the AOE, *i.e.*, $m$-preference as opposed to planarity. In the Bayesian formalism the matter can be addressed by replacing the the covariance matrix (5) by

$$\langle |a_{\ell m}|^2 \rangle (\mathbf{n}) = c_\ell (\delta_{m' |m|} + \epsilon (1 - \delta_{m' |m|})) \qquad (6)$$

where $\mathbf{n}$, $\epsilon$ and $m'$ are the free parameters of the model, with $\epsilon \leqslant 1$. We find that *if we analyze each $\ell$ separately* we rediscover the instabilities reported in Section 2. In Table 3 we take TOHc1 for definiteness, and present the winning $m$, its associated $(b, l)$ and $H_f$; and also the runner up in cases where we get close calls in maximising $H_f$. We see that the Bayesian analysis, in this set up, merely confirms the $\ell = 2$, $m' = 0, 2$ and the $\ell = 4$, $m' = 0, 2$ instabilities.

However, a totally new perspective into these instabilities now makes itself known. $H_f$ only becomes the real evidence $H$ after it is degraded by the penalization $H_p$, related to the number of parameters of the model. If we allow each $\ell$ to choose its own parameters then the overall $H_f$ is large (the sum total) but the penalization is prohibitive as each multipole has 3 parameters. Thus in optimizing $H$ we wish to reduce the number of parameters by always seeking a common axis $\mathbf{n}$ for all $\ell$ in (6). This immediately removes the instabilities found in the frequentist formalism, by effectively penalizing for jumping between close calls, when one choice leads to a better *common* set of parameters.

Take for example $\ell = 2$. We have that $m' = 0, 2$ are close competitors in the optimization of $H_f$; however only $m' = 2$ picks an axis that is roughly aligned with the preferred axis for the other multipoles. So only $m' = 2$ permits a large saving in $H_p$ ($H_p = 2$ per axis, using, say, the AIC) with only small deterioration in $H_f$. An instability would only arise if $m' = 0$ improved $H_f$ by an extra 2 when compared with $m' = 2$. The penalization forces the multipoles to chose common parameters, at the risk of decreasing the fit a little. Thus, in order to maximize $H$—and not only $H_f$—we should select a common $\mathbf{n}$ for $\ell = 2 - 5$, and the complete result (for the same data-set) is presented in Table 4.

In order to mimic the full treatment in Magueijo & Sorkin (2006) we should also seek a common $\epsilon$, thus reducing the number of parameters further. This can be done via the method of Lagrange multipliers, *i.e.*, by maximizing

$$\begin{aligned} H_{Tf} &= \sum_{\ell,i} \frac{N_{\ell i}}{2} \left[ \frac{\sigma_{S\ell i}^2}{\sigma_{\ell i}^2} + \ln \sigma_{\ell i}^2 \right] - \lambda_1 [\sigma_{21}^2 \sigma_{32}^2 - \sigma_{31}^2 \sigma_{22}^2] \\ &- \lambda_2 [\sigma_{31}^2 \sigma_{42}^2 - \sigma_{41}^2 \sigma_{32}^2] - \lambda_3 [\sigma_{41}^2 \sigma_{52}^2 - \sigma_{51}^2 \sigma_{42}^2] \qquad (7) \end{aligned}$$

where $i = 1, 2$ indexes the sub-samples for the $m$-modes with the large and small variance respectively, with $N_{\ell i}$ modes and sample variance $\sigma_{S\ell i}$. The solutions for the variance $\sigma_{\ell i}$ are constrained such that $\sigma_{\ell 2}/\sigma_{\ell 1} = \epsilon$, to fit with our model (6). This has solution

$$\sigma_{\ell i}^2 = \frac{\sigma_{S\ell i}^2}{1 - \frac{2\alpha_{\ell i}}{N_{\ell i}}}$$

with $\alpha_{2i} = \pm A$, $\alpha_{3i} = \pm(-A + B)$, $\alpha_{4i} = \pm(-B + C)$ and $\alpha_{5i} = \mp C$, where $A, B, C$ are solutions of the 3 quadratic equations expressing $\epsilon_2 = \epsilon_3 = \epsilon_4 = \epsilon_5$.

The results are presented in Table 5. For all of the data-sets the choice $m'$s are the same (as opposed to the frequentist statistics), and the preferred common axis is remarkably robust. The common parameter $\epsilon$ and $H_f$ are also reasonably stable. Thus as far as choice of statistics V's available data-sets are concerned we have found an improved formalism.

Regrettably at this point the options for penalization spoil the party, introducing some ambiguities. Undoubtedly using the AIC we find decisive evidence for $m$-preferrence AOE. Bruteforce use of the BIC leads to a very different picture, mainly because the penalization for the axis and $\epsilon$ jumps from $\frac{3}{2} \ln 12 \approx 3.73$ for planarity to $\frac{3}{2} \ln 32 \approx 5.20$.

One may also question the way we treat the penalization of discrete variable $\{m'_\ell\}$. Reflecting on the number of computer bits taken to store it: $(\ln 2 + \ln 2 + \ln 3 + \ln 3)$, makes the

| $\ell$ | $m'$ | $(b$ | $l)$ | $\epsilon$ | $H_f$ |
|---|---|---|---|---|---|
| 2 | 2 | 49 | -96 | 0.052 | 2.33 |
| 3 | 3 | 49 | -96 | 0.108 | 2.01 |
| 4 | 0 | 49 | -96 | 0.058 | 3.21 |
| 5 | 3 | 49 | -96 | 0.028 | 7.34 |
| 2-5 | — | 49 | -96 | —— | 14.89 |

**Table 4.** The evidence $H_f$ for a dominating $m$-mode in the TOHcl map, where each multipole can select its own $\epsilon$ and $m'$, but a common favoured axis is found.

| Data-set | $(b$ | $l)$ | $\epsilon$ | $m$'s | $H_f$ | $H^{AIC}$ | $H^{BIC}$ | $H^{BIC}_{(+)}$ |
|---|---|---|---|---|---|---|---|---|
| LILC1 | 48 | -100 | .077 | 2303 | 11.46 | 4.46 | 2.68 | 4.88 |
| TOHc1 | 49 | -96 | .051 | 2303 | 14.54 | 7.54 | 5.76 | 7.96 |
| TOHw1 | 49 | -96 | .045 | 2303 | 15.25 | 8.25 | 6.47 | 8.67 |
| TOHc3 | 48 | -97 | .073 | 2303 | 11.57 | 4.57 | 2.79 | 4.99 |
| TOHw3 | 48 | -97 | .067 | 2303 | 12.08 | 5.08 | 3.30 | 5.50 |
| WMAP3 | 48 | -100 | .072 | 2303 | 12.10 | 5.10 | 3.32 | 5.52 |

**Table 5.** The optimal $H_f$ for a common axis and common $\epsilon$ between the four multipoles $\ell = 2{-}5$, and the variable $m'$, for various data-sets. We then consider the evidence $H$ using AIC, BIC, and a model with positive mirror parity.

total penalization of 8.78 used in Table 5. If we make it simply the $\frac{1}{2}\ln$ of the number of data points used to determine each: $\frac{3}{2}\ln 32 + \frac{1}{2}(\ln 5 + \ln 7 + \ln 9 + \ln 11)$, then this actually makes the penalization worse (9.27). However if the model has a built in positive mirror parity (de Oliveira-Costa et al. 1996; Starobinsky 1993; Land & Magueijo 2005c), the number of possible $m'$ values is reduced, leading to a lower penalization (6.58) for the same $H_f$ (only mirror positive modes are found), see the last column of Table 5.

## 4 CONCLUSIONS

We have highlighted weaknesses with the original AOE statistic (2) that probed $m$-preference for $\ell = 2{-}5$. These are primarily: 1) lack of robustness: small changes in the data produce very different results, *i.e.*, the statistics are discontinuous; 2) variations with data-set: it is hard to connect varying results to imperfections in the data or the statistic; 3) need simulations to assess significance: no way of penalizing for extra parameters or comparing competing theories, *e.g.*, planarity V's general m-preference.

We have found an improved formalism by employing a Bayesian approach, which cures the instabilities. These were due to the existence of multiple solutions for a given multipole. But bringing in a penalization related to the number of parameters of the model enforces "Occams Razor" and selects solutions where parameters are common between the multipoles.

The Bayesian approach also allows a better assessment of the relative evidence for planarity (correlation between $\ell = 2, 3$, $m' = \ell$ modes) and $m$-preference (a correlation between $\ell = 2 - 5$, $m'$ not restricted). This extends the work of Magueijo & Sorkin (2006) where the low-$\ell$ low-power evidence was assessed, as well as planarity for some data-sets. Comparing Table 2 and Table 5 we see that overall there is more evidence for $m$-preference than for planarity (the exceptions being the BIC for LILC1 and TOHc3), with slightly stronger evidence from the first-year data. We find that although the Bayesian method is more robust, the various information criteria reach different conclusions. The AIC generally provides strong evidence for planarity and decisive evidence for $m$-preference. The BIC provides varying results. This is to be compared to the evidence for scalar spectral index $n_S \neq 1$, where conclusions also differ between IC methodology (Magueijo & Sorkin 2006).

## ACKNOWLEDGEMENTS

We thank the many people who pestered us with the question "is it still there?", and provided useful conversations, most notably Andrew Liddle, Andrew Jaffe, Ofer Lahav, Peter Coles and Carlo Contaldi. Our calculations made use of the HEALPix package (Gorski et al. 1998) and were performed on COSMOS, the UK cosmology supercomputer facility. KRL is funded by the Glasstone fellowship.

## REFERENCES

Bennett C., et al., 2003, Astrophys. J. Suppl., 148, 1
Bernui A., et al., 2006, astro-ph/0601593
Copi C., Huterer D., Schwarz D., Starkman G., 2006, Mon. Not. Roy. Astron. Soc., 367, 79
de Oliveira-Costa A., Smoot G., Starobinsky A., 1996, Astrophys. J., 468, 457
de Oliveira-Costa A., Tegmark M., Zaldarriaga M., Hamilton A., 2004, Phys. Rev., D69, 063516
Donoghue E., Donoghue J., 2005, Phys. Rev., D71, 043002
Eriksen H., Banday A., Gorski K., Lilje P., 2004, Astrophys. J., 612, 633
Eriksen H., et al., 2004a, Astrophys. J., 605, 14
Eriksen H., et al., 2004b, Astrophys. J., 612, 64
Gorski K., Hivon E., Wandelt B., 1998, astro-ph/9812350
Hansen F., Banday A., Gorski K., 2004, astro-ph/0404206
Hansen F., Cabella P., Marinucci D., Vittorio N., 2004, Astrophys. J., 607, L67
Hinshaw G., et al., 2006, astro-ph/0603451
Land K., Magueijo J., 2005a, Mon. Not. Roy. Astron. Soc., 357, 994
Land K., Magueijo J., 2005b, Phys. Rev. Lett., 95, 071301
Land K., Magueijo J., 2005c, Phys. Rev., D72, 101302
Land K., Magueijo J., 2006, Mon. Not. Roy. Astron. Soc., 367, 1714
Liddle A., 2004, Mon. Not. Roy. Astron. Soc., 351, L49
Magueijo J., Sorkin R., 2006, astro-ph/0604410
Park C., 2004, Mon. Not. Roy. Astron. Soc., 349, 313
Ralston J., Jain P., 2004, Int. J. Mod. Phys., D13, 1857
Schwarz D., Starkman G., Huterer D., Copi C., 2004, Phys. Rev. Lett., 93, 221301
Spergel D., et al., 2003, Astrophys. J. Suppl., 148, 175
Spergel D., et al., 2006, astro-ph/0603449
Starobinsky A., 1993, JETP Lett., 57, 622
Tegmark M., de Oliveira-Costa A., Hamilton A., 2003, Phys. Rev., D68, 123523

Vielva P., Martinez-Gonzalez E., Barreiro R., Sanz J., Cayon L., 2004, Astrophys. J., 609, 22